Unit 2- World Wide Web

Introduction to WWW

The World Wide Web (WWW), commonly known as the Web, is an information system where documents and other web resources are identified by Uniform Resource Locators, which may be interlinked by hyperlinks, and are accessible over the Internet.

In simple terms the World Wide Web is a way of exchanging information between computers on the internet and trying them together into a vast collection of interactive multimedia resources.

Elements of WWW

1) **Web Server:**

The resources of the Web are transferred via the Hypertext Transfer Protocol (HTTP), may be accessed by users by a software application called a web browser, and are published by a software application called a web server.

2) *Web pages:*

A web page (also written as webpage) is a document that is suitable for the World Wide Web and web browsers. On a network, a web browser can retrieve a web page from a remote web server. The web browser uses the Hypertext Transfer Protocol (HTTP) to make such requests to the web server.

There are two types of web pages:

1) Static web pages:
   - Pages will remain same until someone changes it manually.
   - In static web pages, Information are change rarely.
2) Dynamic Web pages:
   - Content of pages are different for different visitors.
   - In dynamic web page, Information are change frequently.

3) *Website:*

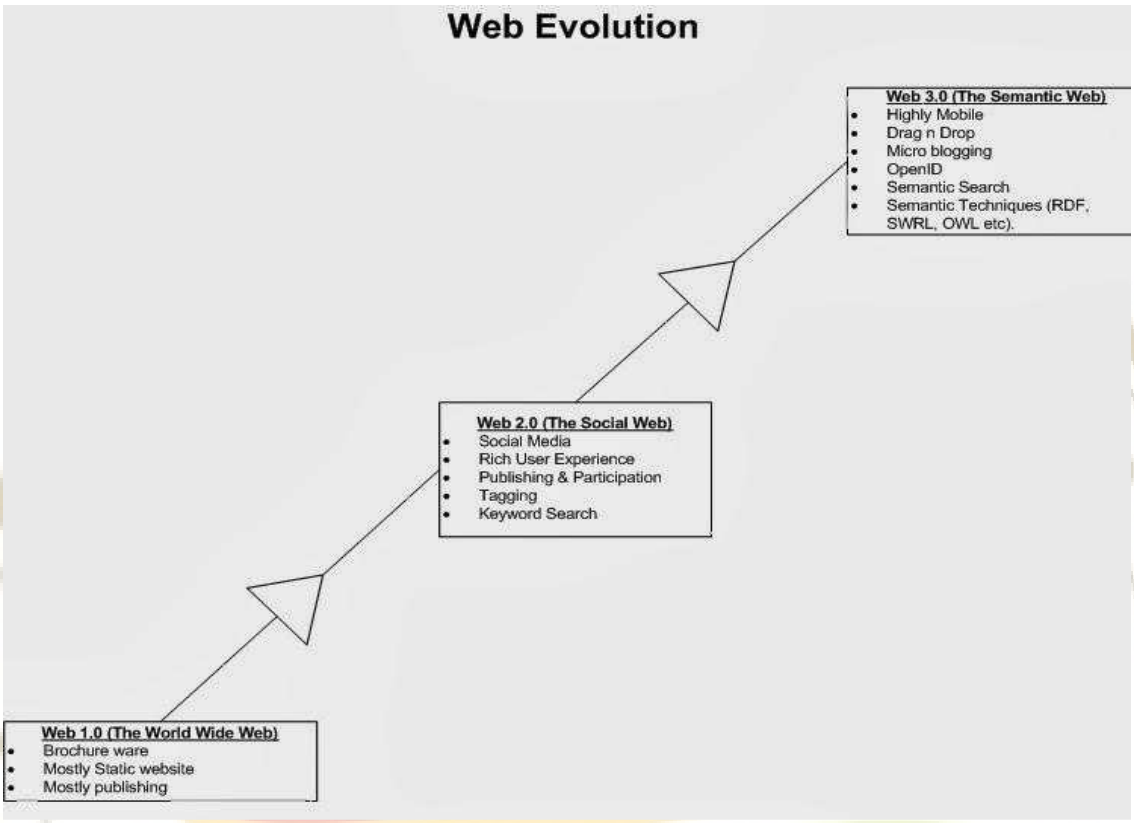
A website is a collection of related web resources including web pages, multimedia content, typically identified with a common domain name, and published on at least one web server.

4) *Web browser:*

A web browser (commonly referred to as a browser) is a software for accessing information on the World Wide Web.

➢ *WWW Evolution*

WWW was created by Timothy Berners Lee in 1989 at CERN in Geneva. The World Wide Web is a system of interlinked hypertext Documents accessed via the Internet. With a web browser, one can view web pages that may contain text, images, videos, and another multimedia and navigate between them via hyperlinks.



- **Characteristics of Web 1.0**

  1) They have read only content. (Static Web page)
  2) Establish an online presence and make their information available to anyone at any time.
  3) It includes static web pages and use basic Hypertext Markup Language.

- **Characteristics of Web 2.0**

  - Technology in center:
    - Web has become a platform with software above the level of a single device. Technology that is associated with blogs, wikis etc.
  - Business in center
    - A way of architecting software and businesses. The business revolution in the computer industry caused by the move to internet.

- User in center
  - The Social Web is often used to characterize sites that consist of people .It is all about content management and new ways of communication and interaction between users. (Ex.Facebook)

➢ **Characteristics of Web 3.0**

- Open Source Software Platform.
- Distributed Database
- Web Personalization.
- Resource Pooling
- Intelligent Web.

Comparison of Web 1.0, Web 2.0 and Web 3.0		
WEB 1.0	WEB 2.0	WEB 3.0
1996 – 2004	2004 -2016	2016+
The Hypertext Web	The Social Web	The Semantic Web
Inventor Tim Berners Lee	Inventor Tim O'Reilly	Inventor Tim Berners Lee
Read Only	Read and Write Web	Executable Web
Millions of User	Billions of User	Trillions+ of Users
Echo System	Participation and Interaction	Understanding self
One Directional	Bi-Directional	Multi-user Virtual environment

Component of Web:

There are 3 components of web:

1) **Uniform Resource Locator (URL):** serves as system for resources on web. It's a address of the website.
2) **HyperText Transfer Protocol (HTTP):** specifies communication of browser and server.
3) **Hyper Text Markup Language (HTML):** defines structure, organization and content of webpage.

WWW Architecture:

1) Client-Server model:

WWW works on client- server approach. Initially, the web consisted of a two-tiered architecture: clients and servers. Clients and servers shared the tasks and services that the system was supposed to perform.

For example, the client may request a service from the server; the server answers the request by providing the service.

Retrieving a website using a URL address that directs to a server to load the site in the client's browser is an example of the two-layer model, also known as the client-server model.

The internet protocol family, which now consists of around 500 different network protocols, is usually used as the basis for the WWW, but it usually comprises the TCP/TCP/IP reference model.

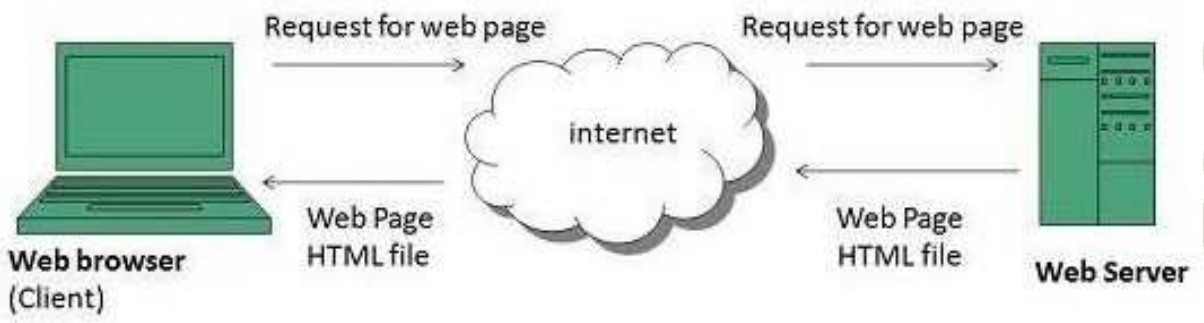
Three prerequisites must exist in the web architecture for the distributed application systems to communicate with one another:

- **Representation formats with a fixed standard:** The most frequently used formats are HTML and CSS; or XML when machines communicate with one another.

- **Protocols for data transfer:** HTTP (Hypertext Transfer Protocol) or HTTPS (Hypertext Transfer Protocol Secure) is used in the web.
- **The standard for addressing:** This refers to the URL (Uniform Resource Locator) which is an instance of the more general concept of URI.

Following steps for how the web works:

- User enters the URL (say, www.svpatelcsbm.org) of the web page in the address bar of web browser.
- Then browser requests the Domain Name Server for the IP address corresponding to www.svpatelcsbm.org.
- After receiving IP address, browser sends the request for web page to the web server using HTTP protocol which specifies the way the browser and web server communicates.
- Then web server receives request using HTTP protocol and checks its search for the requested web page. If found it returns it back to the web browser and close the HTTP connection.
- Now the web browser receives the web page, It interprets it and display the contents of web page in web browser's window.



WWW Client –Server Architecture

Introduction to internet Protocols:

Protocol:

A protocol is a standard set of rules that allow electronic devices to communicate with each other.

Example:

TCP, UDP, IP, FTP, HTTP etc.

1) TCP (Transmission Control Protocol):

- Transmission Control Protocol uses a reliable delivery system to deliver packets to the destination.
- For example Indian Speed Post Service with a certified letter, the receiver must sign for it, indicating the destination actually received the letter: proof of the delivery is provided.
- **TCP** operates under a similar task:
  - It can detect whether or not the destination received a sent packet.
  - With the postal example, if the certified letter got lost, it would be up to you to resend it; with TCP, you don't have to worry about what was or wasn't received—TCP will take care of all the tracking and any necessary resending of lost data for you.
- TCP's main responsibility is to provide a reliable connection and logical service between two devices.
- **TCP** establish a connection before data can be sent. Both the source and destination can simultaneously send data across the session. It uses to check flow control so that a source device doesn't overload a destination with too many packets. It supports data recovery, where any missed or corrupted information can be re-sent by the source. Any packets that arrive out of order, because the segments traveled different paths to reach the destination, can easily be reordered, since packet use sequence numbers to keep track of the ordering.

2) UDP (User Datagram Protocol)

- UDP is a communications protocol that facilitates the exchange of messages between computing devices in a network. It's an alternative to the transmission control protocol (TCP).
- UDP uses a best-effort delivery system, similar to how first class and lower postal services of the Indian Postal Service work. With a first class letter (post card), you place the destination address and put it in your mailbox, and hope that it arrives at the destination.

- With this type of service, nothing guarantees that the letter will actually arrive at the destination, but in most instances, it does. If, however, the letter doesn't arrive at the destination, it's up to you, the letter writer, to resend the letter: the post office isn't going to perform this task for you.
- UDP operates under the same premise: it does not guarantee the delivery of the transport layer segments. While TCP provides a reliable connection, UDP provides an unreliable connection.
- UDP doesn't go through a three-way handshake to set up a connection—it simply begins sending the data. Likewise, UDP doesn't check to see whether sent segments were received by a destination; in other words, it doesn't use an acknowledgment.
- As UDP does not provide assurance of delivery of packet, reliability and other services, the overhead taken to provide these services is reduced in UDP's operation. Thus, UDP provides low overhead and higher speed.

3) FTP (File Transfer Protocol)

- FTP stands for File transfer protocol. FTP is a standard internet protocol provided by TCP/IP used for transmitting the files from one host to another.
- It is mainly used for transferring the web page files from their creator to the computer that acts as a server for other computers on the internet. It is also used for downloading the files to computer from other servers. It provides the sharing of files.
- It is used to encourage the use of remote computers.
- It transfers the data more reliably and efficiently.
- Although transferring files from one system to another is very simple and straightforward, but sometimes it can cause problems. For example, two systems may have different file conventions. Two systems may have different ways to represent text and data. Two systems may have different directory structures. FTP protocol overcomes these problems by establishing two connections between hosts. One connection is used for data transfer, and another connection is used for the control connection.

4) HTTP (HyperText Transfer Protocol)

- HTTP stands for HyperText Transfer Protocol.
- It is a protocol used to access the data on the World Wide Web (www).
- The HTTP protocol can be used to transfer the data in the form of plain text, hypertext, audio, video, and so on.

- This protocol is known as HyperText Transfer Protocol because of its efficiency that allows us to use in a hypertext environment where there are rapid jumps from one document to another document.
- HTTP is similar to the FTP as it also transfers the files from one host to another host. But, HTTP is simpler than FTP as HTTP uses only one connection, i.e., no control connection to transfer the files.

5) IP (Internet Protocol)

- IP stands for internet protocol. It is a protocol defined in the TCP/IP model used for sending the packets from source to destination.
- The main task of IP is to deliver the packets from source to the destination based on the IP addresses available.
- IP defines the packet structure that hides the data which is to be delivered as well as the addressing method that labels the datagram with a source and destination information.
- An IP protocol provides the connectionless service, which is accompanied by two transport protocols, i.e., TCP/IP and UDP/IP, so internet protocol is also known as TCP/IP or UDP/IP.
- The main function of the internet protocol is to provide addressing to the hosts, encapsulating the data into a packet structure, and routing the data from source to the destination across one or more IP networks.

Internet Service Provider (ISP)

Internet Service Provider (ISP) can be defined as "a provider providing home or business users with a connection to access the internet." They can also host web based applications.

- Connecting to the Internet requires specific telecommunications, networking, and routing equipment. ISPs allow users access to networks that contain the required equipment, enabling users to establish Internet connectivity.
- Without an ISP, you wouldn't be able to shop online, access Face book, or read this page.

➢ **Role of ISP**

- As a medium that provides services to connect to the internet.
- Connect customers to the nearest Internet gateway.
- Connecting an information service to a user of the World Wide Web (www).
- Allows a user to use the services of electronic mail (e-mail).
- Allows a user voice conversation via the internet.

- Gave place to the homepage.
- ISP does protection from the spread of the virus by applying antivirus systems for his customers.
- The two main functions of ISP are as follows:

1. **Provide a link:** Firstly, they provide a link to a company or individual who enables them to access World Wide Web and send Internet e-mail. They are the companies that provide individual and institutional subscribers with access to Internet.

2. **Hosting:** Secondly, they host websites or publish a company's website content to enable other companies or consumers access to it. For e.g., a person who is interested in launching a website will first obtain an account with a hosting service provider and then will upload web pages onto his web site which is physically located on the host's 'server'.

Internet Service Provider in India

Sr.No	ISP Name	Users
1	Reliance Jio	388,390,116
2	Airtel	175,680,949
3	Vodafone Idea	139,470,822
4	BSNL	30,907,876
5	ACT Fibernet	1,607,015
6	APSFL	970,270
7	MTNL	1,026,441
8	Excitel	1,350,783

9	Hathway	969,157
10	You Broadband	793,244
11	GTPL Broadband	359,347

Applications of Internet

- Internet Applications can be described as the type of applications that use the internet for operating successfully, that is, by using the internet for fetching, sharing and displaying the information from the respective server systems. It can be accessed only with the help of the internet facility, and it cannot be functional without the internet. There are various internet application are as follow

Search Engine , Web Server, News Group

1) Search Engine

- A search engine is a web-based tool that enables users to locate information on the World Wide Web. Popular examples of search engines are Google, Yahoo!, and MSN Search. Search engines utilize automated software applications that travel along the Web, following links from page to page, site to site.
- Search engines essentially act as filters for the information available on the internet. They allow users to quickly and easily find information that is of genuine interest or value, without go through numerous irrelevant web pages.
- Four main functions of the search engines.

1. **Crawling**

   The crawler, or web spider, is a software component of the search engine. It essentially sorts through the Internet to find website addresses and the contents of a website.

2. **Indexing**

   Once the search engine has crawled the contents of the Internet, it indexes that content based on the occurrence of keyword in each individual website.

3. **Storage**

Storing web content within the database of the search engine is essential for fast and easy searching. The amount of content available to the user is dependent on the amount of storage space available.

4. **Results**

Results are the hyperlinks to websites that show up in the search engine page when a certain keyword is queried. When you type in a search term, the crawler runs through the index and matches what you typed with other keywords.

2) Web Server

- A web server is a computer that runs websites. It's a computer program that distributes web pages as they are requisitioned. The basic objective of the web server is to store, process and deliver web pages to the users.

1. Stores and secures website data:

In web hosting services, a web server stores all website data and secures it from unauthorized users when it is properly configured.

2. Provides web database access:

A web server's responsibility is to provide access to websites that are hosted.

3. Serve the end user requests:

Web servers accept requests from different users connected over the internet and serve them accordingly.

4. Bandwidth controlling to regulate network traffic:

It is a feature available in web server to minimize excess network traffic.

1. News Group

- Newsgroup is an Internet-based discussion group, where people post messages concerning whatever topic around which the group is organized. You can browse newsgroups and post or reply to topics using a newsreader program. Newsgroup names are made up of parts, separated by dots, that indicate the topics covered in the newsgroup.
- An example of a newsgroup is a "computer help" group where individuals assist others with computer problems. Another example is COLA, where Linux information is announced.
- Other example of Newsgroups are as follow

1. alt.best.of.internet
2. alt.binaries.sounds.misc
3. alt.config
4. alt.hackers

E-mail, E-Learning, E-Banking, E-Governance

1) E-mail:

The Internet is commonly used for sending and receiving e-mails. We can send a message electronically to any person on the globe, provided that person has an e-mail-id. This service is fast and economical. For sending or receiving E-main you have an E-mail Address.

An e-mail address (account) consists of two main parts, the username and the domain-server name with the symbol @ in between.

<username>@<domain-name>

Username can be any name but domain-name is fixed for a particular website on which we have our e-mail account. For example, xyz@gmail.com or abc@yahoo.com, etc.

- **Receiving E-mail:**

  To see the mails received, click on Inbox. A number against the Inbox indicates the number of unread mails. Once the Inbox is open, we shall see all the mails listed there. To read a mail, we need to just click on the mail that we wish to read.

  After going through the mails we can either let them remain in our account for future reference or we can also delete the same. There are options for deleting or transferring them to trash. Trash is folder where we keep the mails to be deleted.

- **Sending E-mail:**

  To send an e-mail, both the sender and receiver must have an e-mail account. To send an e-mail, follow the steps listed below:

  1. Click on the 'Compose' option.
  2. Enter the e-mail address of the receiver against any of the three options - To , CC and BCC.
     - ***To:*** Address(es) mentioned in this box are visible to every receiver.
     - ***CC:*** stands for Carbon Copy. The mail will be sent to the To address as well as CC address. The addresses typed in To as well as in CC are visible to all other receivers.

- ***BCC:*** Stands for Blind Carbon Copy. Each user whose addresses are typed in BCC is unaware of the fact that the same message has also been sent to others.

2) E-learning:

E-learning deals with educating people online with the help of the Internet. Development in this field has led to the formation of virtual classes and online sessions. A teacher can teach many students at the same time sitting at different geographical locations. Such classes may be open to all, or may require prior student registration.

E-Learning can be delivered in different ways:

1. ***At the time of learning teacher and students both are to be online.***
   - **Video Lecture :**

     The teacher or instructor can take a class while students can see and listen to him/her. For such a classroom a web camera is required at every terminal.
   - **Textual Lecture :**

     The teacher can take a class by passing his lectures as text and students providing textual response. Students can post their queries online and can get the response instantly.
2. ***At the time of class both teacher and student need not be online at the same time.***
   i) The teacher can send the lecture using tools like explain these and blogs online to the students.
   ii) The students can read their lessons online. If they have any queries then they can mail their queries to the teacher's account.

3) E-banking:

Almost all the nationalized banks have started providing their services online. This means you can have a bank transaction through Internet. E-banking is a facility that allows you to conduct your banking transactions online with laptop or computer with an Internet connection. Every user is allotted a unique user id and password by the bank for this purpose.

The account holder can view account balances and transaction statements, transfer money between accounts in the bank, pay utility bills at any time of the day and from any place, can take printouts of financial-statements, and many more.

4) E-Governance:

E-governance, expands to electronic governance, is the integration of Information and Communication Technology (ICT) in all the processes, with

the aim of enhancing government ability to address the needs of the general public.

The basic purpose of e-governance is to simplify processes for all, i.e. government, citizens, businesses, etc. at National, State and local levels.

E-governance has a great role to play, that improves and supports all tasks performed by the government department and agencies, because it simplifies the task on the one hand and increases the quality of work on the other.

Types of interactions in E-Governance:

1) ***G2G (Government To Government):***

   When the exchange of information and services is within the periphery of the government, is termed as G2G interaction. This can be both horizontal, i.e. among various government entities and vertical, i.e. between national, state and local government entities and within different levels of the entity.

2) ***G2C (Government to Citizen):***

   The interaction amidst the government and general public is G2C interaction. Here an interface is set up between government and citizens, which enables citizens to get access to wide variety of public services. The citizen has the freedom to share their views and grievances on government policies anytime, anywhere.

3) ***G2B (Government to Business):***

   In this case, the e-governance helps the business class to interact with the government. It aims at saving time, cost and establish transparency in the business environment, while interacting with government.

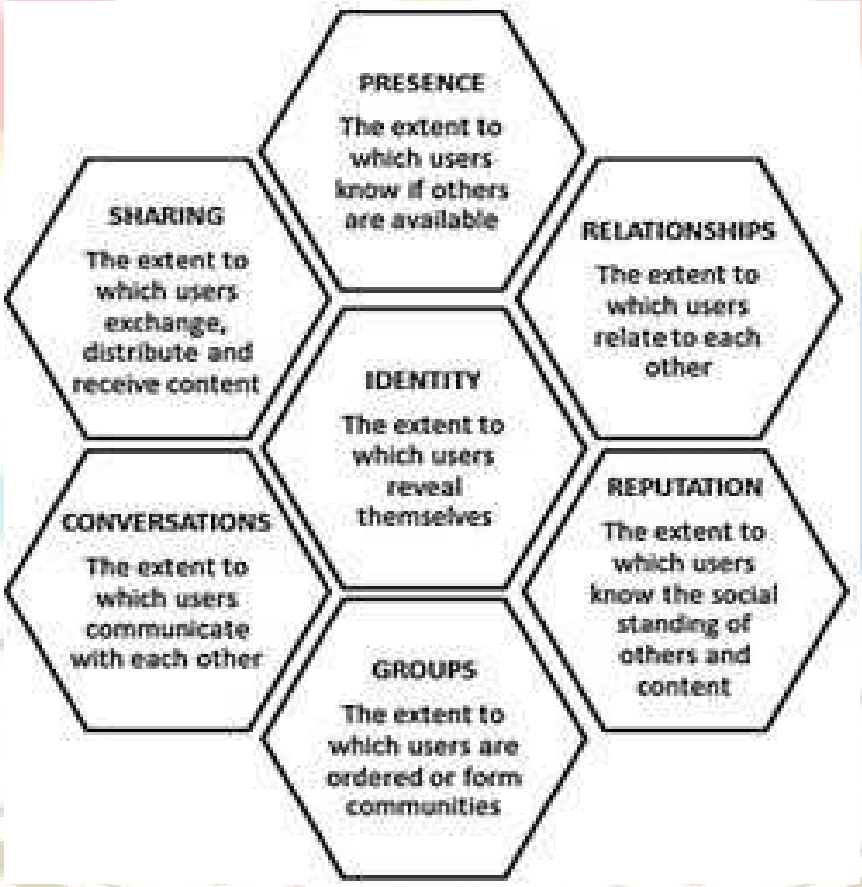
4) ***G2E (Government to Employees):***

   The government of any country is the biggest employer and so it also deals with employees on a regular basis, as other employers do. ICT helps in making the interaction between government and employees fast and efficient, along with raising their level of satisfaction by providing perquisites and add-on benefits.

Social Networking, Instant Massaging, IRC, Audio and Video Conferencing

1) Social Networking

- Social networking is the use of Internet-based social media sites to stay connected with friends, family, colleagues, customers, or clients. Social networking can have a social purpose, a business purpose, or both, through sites like Facebook, Twitter, LinkedIn, and Instagram.
- Social networking sites allow users to share ideas, digital photos and videos, posts, and to inform others about online or real-world activities and events with people in their network.
- A framework that defines social media, which includes seven functional social media building blocks: **identity, conversations, sharing, presence, relationships, reputation, and groups**.



2) Instant Massaging

- Instant messaging (IM), form of text-based communication in which two persons participate in a single conversation over their computers or mobile devices within an Internet-based chatroom.
- One of the core features of many instant messenger clients is the ability to see whether a friend or co-worker is online and connected through the selected service -- a capability known as presence. As the technology has evolved, many IM clients have added support for exchanging more than just text-based messages, allowing actions like file transfers and image sharing within the IM session.
- The exchange of text has long been the main function of instant messaging, but it is now one feature of many. The ability to insert images and emojis into messages is now standard in many clients, as are file transfers. Facebook Messenger even enables users to send money via IM. Numerous clients now support the many from IM to other modes of communication, such as group chat, voice calls or video conferencing.
- Presence enables users to see the availability of their contacts -- not only whether they are online or offline, but also whether they have indicated their status is free or busy. Within an active session between two users, most clients can also indicate to one user in real time when the other user is typing.
- Example of IM are Facebook messenger , Google Talk, Microsoft Skype,Telegram,Viber,WeChat,WhatsApp, Yahoo messenger etc.

3) IRC (Internet Relay Chat)

- IRC (Internet Relay Chat) is a protocol for real-time text messaging between internet-connected computers created in 1988. It is mainly used for group discussion in chat rooms called “channels” although it supports private messages between two users, data transfer, and various server-side and client-side commands.
- Internet Relay Chat is Computer conferencing on the Internet. There are hundreds of IRC channels on numerous subjects that are hosted on IRC servers around the world. “channels” through which you can chat with many people all over world. After joining channel, you can see what other people on this channel type on their keyboards. In that situation, everyone in this channel can see whatever you type on your keyboard.Ex. mIRC, netsplit, SOP and AOP active channels.

4) Audio and Video Conferencing

- Audio conferencing allows people from different locations to interact by telephone through an audio conference bridge. This is also referred to as a conference call.
- Video conferencing is a type of online meeting where two or more people engage in a live audio-visual call. In business, people typically use video conferencing to communicate and collaborate within and outside an organization.
- Modern audio and video technology is saving businesses a significant amount of money every day; companies are now able to save on travel and accommodation expenses.
- A video conference call allows easier decision making for businesses by eliminating the need to convene a physical meeting.
- It saves companies a lot of money by enabling them to reduce office space requirements rather than downsizing their staff. Money is also saved since the need to rent a huge office space is effectively eliminated since most of the staff can work from home.
- The technology allows instant convention of meetings worldwide with little notice.
- Delegates can attend meetings even when they are physically unable to. For instance, in areas where leaders are not allowed to meet due to safety or legal restrictions.
- Audio visual conferencing is an extremely reliable service that operates 24 hours a day. It is also very secure with an encryption that is protected by individual security codes, ensuring that the meeting remains secure and uninterrupted.